DOCUMENT RESUME

ED 102 209 95 TM 004 510

| | |
|---|---|
| AUTHOR | Stodolsky, Susan S. |
| TITLE | Open Education: A Challenge to Evaluators. Report No. 29. |
| INSTITUTION | ERIC Clearinghouse on Tests, Measurement, and Evaluation, Princeton, N.J. |
| SPONS AGENCY | National Inst. of Education (DHEW), Washington, D.C. Office of Dissemination and Resources. |
| REPORT NO | ERIC-TM-29 |
| PUB DATE | Dec 74 |
| CONTRACT | OEC-0-70-3797(519) |
| NOTE | 25p. |
| EDRS PRICE | MF-$0.76 HC-$1.58 PLUS POSTAGE |
| DESCRIPTORS | Affective Objectives; Classroom Environment; Classroom Techniques; Educational Objectives; *Evaluation Needs; *Literature Reviews; *Open Education; *Program Evaluation; Student Evaluation; *Teacher Role; Testing Problems |

ABSTRACT

A review of selected literature on evaluation of open education programs provides a background for the author's discussion of the conditions essential to open education and the evaluation of its effectiveness. Other factors mentioned are: (1) the importance of the types of measuring instruments used to evaluate these programs, (2) the application of the classic evaluation paradigm, (3) lack of procedures for assessing the kinds of behaviors of interest to open education programs, i.e. instruments in the affective domain and classroom observational techniques. (DEP)

ED102209

TM 004 510

BEST COPY AVAILABLE

# Open Education: A Challenge to Evaluators

by

Susan S. Stodolsky

ERIC/TM Report 29

December 1974

[1] The report upon which this article is based was funded by the ERIC Clearinghouse on Tests, Measurement and Evaluation. The Clearinghouse operates under contract with the National Institute of Education, U.S. Department of Health, Education and Welfare. Contractors undertaking such projects under government sponsorship are encouraged to express freely their judgment in professional and technical matters. Points of view or opinions do not therefore represent official National Institute of Education position or policy.

BEST COPY AVAILABLE

# Open Education: A Challenge to Evaluators

Susan S. Stodolsky
University of Chicago

The purpose of this paper is to review selected literature on the evaluation of open education programs. It seems important, as a first step, to clarify what class of educational environment is being discussed when people refer to open education. What is the nature of the beast?

It is not really easy to delimit the meaning of open education, but there have been a number of systematic attempts to define the term and also to empirically describe a classroom which is more open than traditional, the way the teacher might behave, and the sort of child behavior one might see. Some of the most complete and helpful discussions which attempt to define characteristics of open or informal classrooms are those by: Barth (2), Bussis and Chittenden (6); Featherstone (12,13), Flurry (14); Gross and Gross (17), Katz (26); Rogers (28); and Weber (40). Although these authors vary somewhat in the approaches they take, they tend to agree on the central characteristics of open classrooms. They recognize that open classrooms in the United States have been inspired by the informal methods of British teachers (6).

Certain basic assumptions about the way children learn are made by open educators. They accept the developmental view that children

BEST COPY AVAILABLE

learn through direct experience and are curious and motivated to learn. They accept the idea that children differ in the ways they learn and the times at which various aspects of development occur. Open educators believe that children have interests which are expressed in their activity and play and that these tendencies should naturally be respected and nurtured in the school. The developmental perspective adopted by open educators is based on the psychological theories and research of Susan Isaacs [21], Nathan Isaacs [20] and John Dewey.

Beatrice and Ronald Gross [17] present a list of principles which seem to reflect the views in the literature I have seen on open education. They say:

> There are four operating principles of the open classroom. First, the room itself is decentralized: an open, flexible space divided into functional areas, rather than one fixed homogeneous unit. Second, the children are free for much of the time to explore this room, individually or in groups, and to choose their own activities. Third, the environment is rich in learning resources, including plenty of concrete materials, as well as books and other media. Fourth, the teacher and her aids work most of the time with individual children or two or three, rarely ever presenting the same material to the class as a whole (p. 10).

The notion of the "integrated day" is an important one in the British literature on informal education. The core idea is that subject matters and contents of learning are not dealt with singly but that the child's behavior and activity can lead to the use of various subject matters in a given activity. Learning through the development of an interest in countries, colour tiles by grade from page 10 to page 11, is favored by many practitioners of behavioral methods

[redacted]

BEST COPY AVAILABLE

and renewed efforts hould be supported, I have some recurrent doubts about the potential of such efforts. The use of criterion-referenced rather than norm-referenced instruments may be helpful (15, 37). Yet my work with children in different kinds of classrooms makes me question whether the psychometric approach will prove relevant to many of the constructs which interest us. I find it somewhat consoling that other colleagues involved in the evaluation of educational programs have voiced similar doubts. Herb Zimiles (42), in a discussion of personality measurement, suggests that Allport's idiographic view may have merited more support by psychologists than it received.

> . . . different traits are differentially salient for different children (or adults). Across-the-board measurement of a particular trait generates a hodge-podge of data. The data gathered from those for whom the trait is salient may be quite telling, but a large portion of the data obtained from the rest of the sample may have no functional significance (p. 5).

Recently, I have been attempting to understand the choices children made in classrooms which were moving toward an open philosophy. I used a series of psychometric instruments which I thought might help differentiate children in conjunction with observations of children's classroom behavior. I was interested in finding out what sorts of children select what sorts of activities and exhibit different styles in an open setting. There appear to be relationships between some of the characteristics I measured (intelligence, cognitive style, locus of control, and associational fluency) and certain aspects of the child's functioning in the classroom--for example, the extent to which he pursues activities alone. However, various aspects of child

BEST COPY AVAILABLE

functioning seem to relate to somewhat independent clusterings of child variables (and, of course, some relate to none of the variables). Furthermore, our clinical impressions of the children lead us to very different hypotheses about the correlates and etiology of similar behavior patterns in different children.

What I am trying to illustrate here is the perennial tension between the recognition of individuality and the attempt of measurement to provide comparability of data under standard conditions. While both endeavors are necessary, we may be least successful in applying the psychometric approach in the socioemotional area. Nor am I certain that a quantitative view is appropriate for many aspects of personality and affect. Many of the constructs which interest open educators are more qualitative in nature and depend upon certain conditions for their expression. For example, I am not sure that it is meaningful to ask whether a child has a more positive or negative self concept than another child rather than knowing the nature of his views of himself. Similarly a construct like initiative seems to come in many behavioral forms which are difficult to assess in a uniform manner across children.

An extremely insightful article by Edna Shapiro (31) contains a discussion of the relative utility of tests and classroom observations as sources of data. She says:

> The rationale of the test, . . . is that each child is removed from the classroom and treated equivalently, and differences in response are presumed to indicate differences in what has been taken in, made one's own, that survives the shift to a different situation.

BEST COPY AVAILABLE

> . . . This requires reconsideration of the role of classroom data, individual test situation data, and the relation between them. If we minimize the importance of the child's behavior in the classroom because it is influenced by situational variables, do we not have to apply the same logic to the child's responses in the test situation, which is also influenced by situational variables (pp. 533-34)?

Shapiro does not consider classroom data and test data to be equally useful for all evaluation purposes, but she seriously questions whether the sort of test data we are able to obtain will be any more useful for program evaluation than classroom observations would be. She eventually argues that evaluation should play its chief role as a contributor to program development and as such would justify more flexible uses of observations and experimental instruments. While I agree with this general point, I am also sympathetic to policy makers who want evaluation data in order to make policy decisions about the effectiveness of programs. Perhaps a consideration of some suggestions and studies made by proponents of open education and evaluators who do not hold the classic view will assist us to reach some useful conclusions about the best evaluation strategies.

Alternative Views of Evaluation

A number of persons involved in the field of evaluation have questioned the appropriateness of measuring outcomes as the basis for judging program effectiveness. Many open educators advocate a process of documentation of child growth and development based on children's work and actual classroom behavior (see Carini, 7, Chittenden and Bussis, 9, and Dean, 10). Some helpful examples are presented in a recent article by Hawes (18) which reports the work of Ruth Aldrich at Marcy School, Lillian Weber, Mary Stitt and others. A number of

BEST COPY AVAILABLE

evaluators, including myself (31, 33, 35), have also recognized that the classroom behavior of children be considered valid and important evidence about a program.

There are several positive aspects to this general strategy. First, it is important to better understand the actual classroom experiences and processes of children in diverse educational settings. Data which details the salient features of various classroom environments and of different children's experiences in different classrooms can be used for a number of highly significant purposes.

Most germane here is the use of such descriptive data to make judgments about the quality of an educational program. Zimiles (32) persuasively argues for:

> . . . systematic and comprehensive evaluation of the child's school environment, to be followed by a theoretical analysis of the potential impact of his school experience.
> . . . It is a paradox that we have the responsibility and the capacity comprehensively to describe and record the essential character of an educational program, yet do not do so; and at the same time, we do not know how to assess the impact of a complex set of experiences on the psychological functioning of a developing child, yet we persist in trying to do so (pp. 7-8).

Zimiles believes that we can look at educational environments and make value judgments about them. Such judgments would be based on our theories about the educational conditions which lead to certain desirable ends. Zimiles stresses the necessity for a conceptual framework which would allow us to make connections between what we see in classrooms and our ideas about the "potential impact" of these

BEST COPY AVAILABLE

conditions on the child. I believe that heads of English schools and inspectors who routinely evaluate and advise British teachers implicitly use this strategy.

A related point of view is presented by Chittenden, Bussis, and Amarel at ETS (see Hawes, 18). They suggest that the quality of the classroom experience should form the basis for evaluation of open classrooms. They would evaluate diverse aspects of both student and teacher behavior and products as well as aspects of the physical environment of a classroom. In contrast to conventional evaluation which judges student responses to some uniform condition, they are trying to develop standards for judging diverse student productions and activities. The ETS approach to evaluation focuses on standards of quality regarding the "process, content and context of the child's learning" (Hawes, 18).

Both the ETS group and Zielles endorse the need for evaluation evidence to judge programs. They are searching for new ways of producing evidence which will in some sense be objective and comparable across educational environments. I share their concern about data which will allow one to make judgments about different educational experiences, but I believe that program evaluation per se may be less important for the educational enterprise than evaluation which will feed back into student learning and studies which will begin to explicate the nature of the empirical relationships between various educational experiences, child characteristics, and child outcomes.

BEST COPY AVAILABLE

It may not be necessary to pit one type of work against the other, but I believe large scale cluster evaluations keep shifting too much of our limited resources without adequately answering the key questions.

Open education presents a real challenge to evaluators and educational researchers because it broadens the scope of educational environments and experiences to be understood. We have begun to develop methods for systematically describing some of the relevant features of classroom structure and process. Promising observational procedures have been developed in response to the need for systematic descriptions of educational environments (1, 14, 23, 24, 31, 37).

There appear to be some promising starts on the development of an empirical base for understanding how educational environments differ (4, 16, 32, 36). Some possible consequences of the educational differences are also being studied. For example, in a recent paper using Follow Through data, Stallings (32) related differences in classroom conditions to differences in children's classroom behaviors and to tested outcomes. With respect to open classroom, she says:

> In the more open, interdisciplinary classroom, where a wide variety of activities are occurring, a wide variety of materials are available, children can select their own groupings part of the time, and can engage in activities without adults, children have higher scores on the Raven's perceptual problem-solving test. They are also absent less often, and they take more responsibility for their coming as measured on the Intellectual Responsibility Scale. They are less negativistic, cooperate more often, and ask more questions (p. 11).

facilitative and subjective of this sort of work should be a high priority.

My students and I have been attempting to contribute in the area by behavioral audit-style studies which fullblock naturalistic children such time in free-choice classrooms. We have focused also on the settings of children's behavior by action variables in the children's experiences which occur in a given classroom. These variables in activity type, in turn, have related very consistently to differences in complex outcome measures (see Barlow, 1971). Since students clearly do pursue unique paths within an open classroom, it seems educationally important to know more about the types of choices they make and the types of behaviors they exhibit. In regard to accurate possible correlates of choices and styles both in terms of individual differences in the students and in relation to the situational determinants of choice. Watch [XX] has been able to show in the nursery school that identifiable style differences are strongly predictive of differences in the free play behavior of children, particularly with respect to the pacing of activity and patterns of orientation. My most recent work with seven- and eight-year-olds also seems to be revealing that some systematic variation in child behavior in the open environments is related to meaningful individual differences.

In addition to process-oriented studies, we should permit the evaluations of program innovations by evaluations of students learning and development via feedback systems. Teachers continuously diagnose

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

During the writing of this paper, I have felt frustrated with the task. I strongly feel the lack of highly specific, concrete suggestions on what I have written. Hopefully the many general considerations with which scholarship is always more intimately [illegible] the day-to-day process of preparing to the [illegible] I have tried to provide enough information a way to begin building articles and [illegible] which are currently questions in print. Sadly, the finish now and could only for you. Dear Barbara, the usefulness of such should not overcome my knowledge and [illegible] feelings for [illegible] which [illegible] both physics and [illegible] in

1. Anderson, J. Some new species from Africa. J. Bot. Soc. Trop. Africa 12: 1-15, 1902.

2. Smith, W.G. Some observations on the development of the embryo. Ann. Bot. 10: 1-20, 1896.

3. Garfield, W.C., Hoffman, T., Ellis-Jones, R.T., & Wilson, C.A. A synthetic approach to the study of disorders and disease. In Medical-physical sciences. Johnson Foundation Symposia, Vol. 14, p. 1-21.

4. Bennett, J. The experimental study of behavioral changes in the Drosophilidae. Journal of Comp. & Physiol. Psych. Monograph Supplement, 1939.

5. Bradley, R.S. Variations of temperature in mountains. Quart. Journal Royal Meteorol. Society 43: 1-24, 1917. (See also Bradley & Fahey and Bradley & Fahey, 1918.)

6. Bradley, R.S., and Hitchcock, A.E. Climatic relationships and thermal data. U.S.D.A. Weather Bureau Bull. 1918. Washington: U.S. Government Printing Office, 1919. Pp. 1-24.

7. Carter, R.E. Photoperiodism in relation to hibernation in the bat. Proc. Pan American Congress for Bat Investigations, May 1939, 7: 1-20.

8. Central Statistical Board. The Canadian Statistical Review: Monthly publication. Ottawa: Dominion Bureau of Statistics and Statistical Branch, 1940.

9. Chamberlain, T.C., & Salisbury, E.W. Some preliminaries concerning soil organisms. U.S.D.A. Weather Bureau Bull. 24. Report of the Work of the Agricultural & Soil Chemist. Pp. 35-79. Washington: Government Printing Office, 1902.

10. Craig, R.D., et al. Observations on environmental stress and survival of the Craig Mountain porcupine (Erethizon dorsatum craigi). Journal of Mammal. Monograph. Supplement, Vol. 1, 1940. Pp. 1-274.

11. Davis, R. Influenced cycles in Malthusian terms. Proceedings of Fisher's symposium. New York: Columbia University Press, 1939.